Specification page 26, lines 17-25. No new matter is introduced by entry of the subject amendments.

II. The Double Patenting Rejection

Claims 40-45 are rejection under the judicially created doctrine of obviousness-type double patenting over claims 1-27 of U.S. Patent No. 5,714,596. Applicants respectfully traverse this rejection.

It is admitted by the Examiner that claims 40-45 of the present application are of different scope than the claims of the '596 patent. There is no motivation in the claims of the '596 patent to add a pharmaceutically acceptable excipient, a required element of each claims 40-45, to the oligonucleotide. Thus, the obviousness-type double patenting rejection is improper. Accordingly, applicants request reconsideration and withdrawal of the double patenting rejection.

III. The rejection under 35 U.S.C. 112, First Paragraph

Claims 40-48 and 52-55 were rejected under 35 U.S.C. 112, first paragraph, as containing subject matter which was not described in the specification in such a way so as to enable one skilled in the art to make or use the invention. The thrust of the rejection appears to be that applicants have not provided any guidance as to which portion of the genome would be useful as antisense polynucleotides, nor does the Specification disclose any particular working example. Applicants respectfully traverse the rejection and its supporting remarks.

The Examiner appears to be requiring Applicants to provide working examples of the claimed antisense compositions. Under controlling precedent, however, working examples are not required

in order to enable an invention. *In re Long*, 151 U.S.P.Q. 640, 642 (C.C.P.A. 1966). Whether the specification does or does not contain working examples is only one factor to e considered in determining enablement. *In re Honn*, 150 U.S.P.Q. 652, 657 (C.C.P.A. 1966). The relative skill of those in the relevant art must also be considered. *Ex parte Forman*, 230 U.S.P.Q. 546, 547 (Bd. Pat. App. Interf. 1986).

The present specification is addressed to those of skill in the art, and, the law is clear that the specification need not provide knowledge which is generally known by those of skill in the art. Applicants can properly rely on common knowledge in the art to bolster and supplement its disclosure. *Hybritech Inc. v. Monoclonal Antibodies, Inc.*, 231 U.S.P.Q. 81, 94 (Fed. Cir. 1986); *Genentech Inc. v. Novo Nordisk A/S*, 42 U.S.P.Q.2d 1001, 1005 (Fed. Cir. 1997). The present specification, therefor, need only "supply the novel aspects of [the] invention in order to constitute adequate enablement." 42 U.S.P.Q.2d at 1005.

As mentioned in the response filed on June 16, 1998, the remarks of which are incorporated herein, the specification provides guidance for the design of antisense polynucleotides. See page 78, beginning at line 26, which teaches which portion of the HCV genome the antisense molecules should target. Section II.H., beginning at page 61 of the Specification describes the design of the antisense oligonucleotides and Section II.L of the Specification describes assays for screening techniques for antisense polynucleotides. In addition, the Specification provides the sequence on an HCV genome, HCV1, as well as the information necessary for obtaining other HCV isolates. Thus, one of skill in the art is enabled by the present Specification to make or use the invention by using the assays described in the Specification

to select the antisense oligonucleotides most appropriate for use.

The Office Action appears to assert that the present Specification does not provide evidence that Applicants had made the invention because "at no point in the file history of this application are specific antisense polynucleotides set forth". *Amgen Inc. v, Chugai Pharmaceutical Co. Ltd*, 18 U.S.P.Q.2d 1016 (Fed. Cir. 1991) and *Fiers v. Revel*, 25 U.S.P.Q.2d 1601 (Fed. Cir. 1993) are cited to support this assertion. These cases are misapplied to the present specification. As is admitted in the Office Action, Applicants have disclosed the genomic sequence of HCV1. Applicants submit that the antisense molecules are described because each of the 8, 10, 12, 15 or 20mer HCV1 nucleotides that are candidates for the claimed antisense compositions can be written down and are fully supported in the specification. Thus, applicants have not only "distinguished the material from other materials", and are not in the position of "simply wishing to know the identify of any material with that biological property", they have also defined the candidate molecules by "something other than its function utility." See *Amgen* and *Fiers*. Armed with such a list of candidate molecules, it is then routine experimentation to screen the candidate molecules for antisense activity.

Further, the specification provides all of the information necessary for one of ordinary skill in the art to find other isolates of HCV. Indeed, the sequence of two Japanese isolates of HCV were determined by Okamoto et al in 1990, see attached article, using the information disclosed in European Patent Publication 318,216A1, one of the published European applications related to the present specification. The '216 publication is

specifically acknowledged in the abstract and second paragraph of Okamoto et al as providing the sequence information used to isolate sequences from the Japanese isolates. The sequence data were obtained using primer extension and cloning methods that were known in the art. Once one of ordinary skill has the sequence of different HCV isolates, one could then, following the information in the present specification, identify the candidate molecules and screen them for activity. Simply because it may not be possible for one to predict, *a priori*, which of the candidate molecules would be useful does not render the Specification nonenabling, if screening the candidate molecules merely involves routine experimentation.

As above-demonstrated, *Amgen* and *Fiers* are distinguishable from the present application. Under controlling precedent, the disclosure in the present specification is more than adequate to satisfy the enablement requirement of 35 U.S.C. 112, first paragraph. Accordingly applicants respectfully request reconsideration and withdrawal of the rejection.

Conclusion

In light of the foregoing amendments and remarks, it is believed that the application is in condition for allowance. Accordingly, reconsideration and favorable action on all claims is earnestly solicited. If there are any questions concerning this communication, the Examiner is invited to call the undersigned at the telephone number provided below so that prompt disposition of this application can be achieved.

The Assistant Commissioner is hereby authorized to charge any additional fees (or credit any overpayment) associated with this communication and which may be required under 37 CFR 1.16 and 1.17 to Deposit Account No. 03-1664. This, however, is not authorization to pay the issue fee.

Respectfully submitted,

Dated: 6 March 1996

By: _____
Alisa A. Harbin
Reg. No. 33,895

CHIRON CORPORATION
Intellectual Property - R440
P.O. Box 8097
Emeryville, CA 94608
Telephone: (510) 923-2708

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

[Japan. J. Exp. Med. Vol. 60, 3, p. 167–177, 1990]

The 5′-Terminal Sequence of the Hepatitis C Virus Genome

(Received for Publication, March 2, 1990)

Hiroaki OKAMOTO, Shunichi OKADA*1, Yasushi SUGIYAMA, Shigeru YOTSUMOTO, Takeshi TANAKA*2, Hiroshi YOSHIZAWA*3, Fumio TSUDA*4, Yuzo MIYAKAWA*5 and Makoto MAYUMI[1]

*Immunology Division, Jichi Medical School, Tochigi 329-04, *1First Department of Internal Medicine, Yamanashi Medical College, Yamanashi 409-38, *2Japanese Red Cross Blood Center, Saitama 362, *3Department of Public Health, Hamamatsu University School of Medicine, Shizuoka 431-31, *4Section of Immunology, Kitasato Institute, Tokyo 108, and *5Mita Institute, Tokyo 108, Japan*

Summary: The 5′-terminal sequence of the genome of hepatitis C virus (HCV) was determined for two distinct HCV strains in human and chimpanzee carriers. It had a 5′-noncoding region of at least 324 nucleotides, well preserved by the two strains with a high homology (99.1%), followed by 1348 nucleotides that continued to the documented sequence of prototype HCV spanning 7310 nucleotides (European Patent Application #88310922.5). Based on these results, HCV is considered to possess an uninterrupted open reading frame encoding at least 2886 amino acid residues. Two structural genes were postulated on the 5′-terminal sequence of the HCV genome. One gene in the upstream region, highly conserved by the two strains at the amino acid level and rich in basic amino acids such as arginine, appeared to encode the viral capsid protein. The other gene in the downstream region was divergent between the two strains at both nucleotide and amino acid levels. It coded for nine potential N-glycosylation sites, and was considered to encode the viral envelope protein. Disclosure of the 5′-terminal sequence of the HCV genome would facilitate its taxonomic classification, and contribute toward immunological diagnosis of infection and development of vaccines.

Exclusion of blood units contaminated with hepatitis B virus, by the detection of hepatitis B surface antigen, has decreased the incidence of posttransfusion hepatitis substantially [1]. This, however, has brought to the fore non-A, non-B (NANB) hepatitis, which accounts for most cases of posttransfusion hepatitis at present [2, 3]. Despite recognition of NANB hepatitis in early 70's [4, 5] and transmission studies in chimpanzees that followed [6, 7], the causative virus has remained elusive.

Recently, Choo *et al.* [8] propagated a cDNA clone of the etiological agent, now designated hepatitis C virus (HCV), by probing with antibodies possibly directed to the virus, and identified an RNA virus with a genome size of approximately 10 kilobases. The nucleotide sequence is documented for nearly three-quarters of the HCV genome toward its 3′-terminus (European Patent Application #88310922.5). Antibodies directed to HCV are detectable by immunoassay involving a fusion protein expressed in yeast harboring a part of this sequence [9], and are useful in various clinical and epidemiological settings [10–14].

The nucleotide sequence of HCV disclosed so far (European Patent Application #88310922.5) represents nonstructural genes; the nucleotide sequence of the structural genes of HCV has not been documented at this moment. With the advent of an oligonucleotide primer copied from the nonstructural sequence

[1] To whom correspondence should be addressed.
岡本宏明, 杉山 靖, 四元 茂, 真弓 忠 (自治医科大学予防生態学教室)
岡田俊一 (山梨医科大学第一内科)
田中建志 (埼玉県赤十字血液センター研究部)
吉沢浩司 (浜松医科大学公衆衛生学教室)
津田文男 (北里研究所免疫研究室)
宮川侑三 (三田研究所)

of HCV, we obtained, from cDNA libraries of human and chimpanzee carriers, clones that possessed an apparent 5′-terminal sequence of the HCV genome. The 5′-terminal nucleotide sequence, including a noncoding region of 324 nucleotides, was determined for two HCV strains from human and chimpanzee carriers, respectively, and compared against each other for portraying the structural genes of HCV.

Materials and Methods

Plasmas containing HCV: Four plasmas were selected as the source for propagating cDNA clones of HCV genomes. Two were from blood donors (HC-J1 and HC-J3), and positive for antibody to HCV (anti-HCV) determined by a commercial assay kit (Ortho HCV Ab ELISA Test, Ortho Diagnostic Systems Tokyo, Japan). The other two was confirmed for the infectivity for NANB

Table 1 Nucleotide sequences of 20-mer oligonucleotide primers used for PCR on cDNA of HCV genomes in the four plasmas[a].

Primers	Nucleotide positions	Polarity	Sequence (5′—3′)
#1	37-56	(+)	TTACCGATTTTGACCAGGGC
#2	67-86	(+)	TCAGTTATGCCAACGGAAGC
#3	126-145	(+)	AAACCTTGCGGTATTGTGCC
#4	153-172	(+)	AGTGTGTGTGGTCCGGTATA
#5	268-287	(−)	CGGTGGCCTGGTATTGTTAA
#6	303-322	(−)	GAGTTCATCCAGGTACAACC
#7	329-348	(−)	CTCCGCACACTTTGGTGAAT
#8	401-420	(−)	GATGCTTGCGGAAGCAATCA
#9	6427-6446	(+)	AGATGGCTTTGTACGACGTG
#10	6490-6509	(+)	TCCAATACTCACCAGGACAG
#11	6761-6780	(−)	CACAGCTAGTTGTCAGTACG
#12	6786-6805	(−)	TTGATGTAGCAAGTGAGGGT
#16	4029-4048	(−)	CTGGTGACAGCAGCTGTAAA
#17	4061-4080	(−)	TGAAGAGGAGGGTTTGGCTA
#21	3669-3688	(+)	TATTGCCTGTCAACAGGCTG
#22	3759-3778	(+)	CGAGAGTTCGATGAGATGGA

[a]: Copied from and numbered according to the documented sequence of prototype HCV (European Patent Application #88310922.5).

Table 2 Nucleotide sequences of 20-mer oligonucleotide primers used for cDNA synthesis and PCR.

Primers	Nucleotide positions	Polarity	Sequence (5′—3′)
#8[a]	2073-2092	(−)	GATGCTTGCGGAAGCAATCA
#23[b]	450-469	(+)	TAGATTGGGTGTGCGCGCGA
#25[b]	807-826	(−)	TCCCTGTTGCATAGTTCACG
#30[b]	721-740	(+)	CTCATGGGGTACATTCCGCT
#42[c]	1864-1883	(−)	TCGGTCGTCCCCACCACAAC
#44[d]	1-20	(+)	GGCGACACTCCACCATAGAT

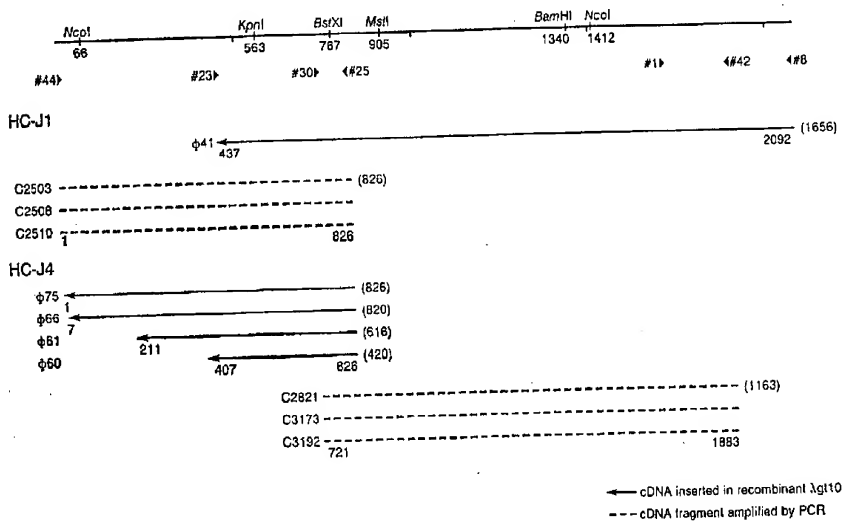
[a]: Copied from the documented sequence of prototype HCV (European Patent Application #88310922.5), and numbered by a new system starting from the putative 5′-terminus of HCV genome (Fig. 1).
[b]: Copied from the sequence of ϕ41 from HC-J1.
[c]: Copied from the sequence of C2127 from HC-J3.
[d]: Copied from the sequence of ϕ75 from HC-J4.

hepatitis by chimpanzee transmission experiments. One of them (HC-J2) was from a patient with chronic NANB hepatitis and positive for anti-HCV, while the other from a chimpanzee (HC-J4) was negative for anti-HCV.

Preparation of nucleic acids and cDNA synthesis: Plasma (1.8 ml) was overlaid on 1 ml of Tris-HCl buffer (10 mM, pH 8.0) containing 20% (w/v) sucrose and 1 mM EDTA, and centrifuged in a Beckman TLA 100.3 angle rotor (Beckman Instruments, Inc., California, U.S.A.) at 68×10^3 rpm for 1 hr. The pellet was suspended in 200 μl of Tris-HCl buffer (50 mM, pH 8.0) containing 200 mM NaCl, 10 mM EDTA, 2% (w/v) sodium dodecyl sulfate and 1 mg/ml proteinase K, and incubated at 60°C for 1 hr. Nucleic acids in the solution were extracted with phenol/chloroform, precipitated with ethanol and then dissolved in 20 μl of distilled water.

Oligonucleotide primers (20-mers), copying the sequences of the prototype HCV (European Patent Application #88310922.5) and HCV genomes in HC-J1, HC-J3 and HC-J4 were synthesized on a 380B DNA synthesizer (Applied Biosystems Japan, Tokyo, Japan). RNA extracted from plasmas was denatured at 70°C for 1 min, and served as a template in the synthesis of cDNA with use of 10 units of

Fig. 1. Strategy for determining the 5′-terminal sequences of two HCV genomes.



The 5′-terminal sequence with restriction sites, and positions of primers used for cDNA synthesis and PCR are indicated above. Nucleotides were numbered by a new system starting from the putative 5′-terminus of HCV genome. cDNA clones and PCR products, for sequencing HCV genomes in HC-J1 and HC-J4 plasmas, are given below. For sequencing the HCV genome in HC-J1, cDNA clone (ϕ41) was obtained from a library primed with #8 that copied a sequence of the prototype HCV (European Patent Application #88310922.5); the sequence after nt 1673 is disclosed in the patent. Based on the upstream sequence of ϕ41, primer #25 (nt 807–826) was synthesized and used to construct a cDNA library on HC-J4; HC-J1 failed to develop a cDNA library with primer #25. The 5′- end of the longest cDNA clone from HC-J4 (ϕ75) was tentatively taken as the start of the 5′-terminus of HCV, and used as the origin in a new numbering system. PCR was performed on cDNA from HC-J1 with primers #44 and #25, and on cDNA from HC-J4 with primers #30 and #42. The entire 5′-terminal sequence of HCV strains of HC-J1 and HC-J4, in this manner, was complemented by cDNA clones from λgt10 libraries and those amplified by PCR.

reverse transcriptase (cDNA Synthesis System Plus, Amersham Japan, Tokyo, Japan) and 20 pmol of oligonucleotide primers. The sequences of 20-mer oligonucleotide primers are listed in TABLES 1 and 2.

Amplification of cDNA with polymerase chain reaction (PCR): PCR was carried out in a DNA thermal cycler (Perkin-Elmer Cetus, Connecticut, U.S.A.) for 35 cycles with a Gene Amp DNA amplification reagent kit (Perkin-Elmer Cetus) in accordance with the method originally described by Saiki *et al.* [15]. Each reaction cycle included denaturation at 94°C for 1 min, primer annealing at 55°C for 1.5 min and primer extension at 72°C for 3 min. PCR products were separated on a composite gel of 1–2% NuSieve and 1–2% SeaKem agarose (FMC BioProducts, Maryland, U.S.A.) and visualized under ultraviolet light after staining with ethidium bromide.

Construction of cDNA libraries: RNA extracted from 10–20 ml of two plasmas (HC-J1 and HC-J4) was employed to construct cDNA libraries in the bacteriophage λgt10 (Amersham). Synthetic oligonucleotides (#8 and #25) were used for priming cDNA synthesis. Recombinant λgt10 phage thus prepared was probed by plaque hybridization with DNA fragments labeled with [^{32}P]. A DNA fragment of 384 nucleotides, amplified by PCR with

TABLE 3 Nucleotide sequences amplifiable on cDNA from the four plasmas by PCR with various combinations of primers.

cDNA primed with	PCR on cDNA primed with	Products of PCR[a] (nucleotides)	Plasmas containing HCV			
			HC-J1	HC-J2	HC-J3	HC-J4
Region A (nt 37-420)						
#5	#1/#5	251	+	−	−	−
#5	#2/#5	221	+	−	−	−
#5	#3/#5	162	+	−	−	−
#5	#4/#5	135	+	+	+	−
#6	#1/#6	286	+	−	−	−
#6	#2/#6	256	+	−	−	−
#6	#3/#6	197	+	−	−	−
#6	#4/#6	170	+	+	−	−
#7	#1/#7	312	+	−	−	−
#7	#2/#7	282	+	−	−	−
#7	#3/#7	223	+	−	−	−
#7	#4/#7	196	+	−	−	−
#8	#1/#8	384	+	−	−	−
#8	#2/#8	354	+	−	−	−
#8	#3/#8	295	+	−	−	−
#8	#4/#8	268	+	−	−	−
Region B (nt 3669-4080)						
#16	#21/#16	380	+	−	−	−
#16	#22/#16	290	+	+	+	+
#17	#21/#17	412	+	+	−	−
#17	#22/#17	322	+	−	−	−
Region C (nt 6427-6805)						
#11	#9/#11	354	+	+	+	+
#11	#10/#11	291	+	−	−	−
#12	#9/#12	379	+	+	+	+
#12	#10/#12	316	+	−	−	−

[a]: PCR products amplified were electrophoresed and detected by staining with ethidium bromide.

primers #1 and #8, was used for screening the cDNA library primed with #8. For screening the cDNA library primed with #25. The other DNA fragment (377 nucleotides) was employed which had been amplified with #23 and #25 (Fig. 1).

Cloning and sequencing: cDNA fragments amplified by PCR were visualized by staining with ethidium bromide. A portion of gel was cut out, corresponding to the position of expected PCR products. DNA eluted from it was treated with T4 polynucleotide kinase (Takara Biochemicals, Kyoto, Japan). The 5'-phosphorylated PCR products, or cDNA inserts excised from λgt10 phage vector, were cloned into the *Hin*cII site or the *Eco*RI site of M13 mp18 phage vector. Three separate clones were isolated from each PCR product, and nucleotide sequences of both plus and minus strands were determined by dideoxy-chain termination method [16].

RESULTS

Genomic heterogeneity among HCV strains in the four plasmas.

Amplification of 24 target sequences of HCV genomes in the four plasmas was attempted by PCR with use of 16 oligonucleotide primers copied from the prototype HCV (European Patent Application #88310922.5). The results are given in TABLE 3. A wide discrepancy was observed in detecting 24 target sequences among HCV strains in the four plasmas. All 24 sequences were amplified in HC-J1, in a fidelity much higher than only 6 in HC-J2, 4 in HC-J3 and 3 in HC-J4. Inasmuch as an inefficient amplification was attributable to the mismatch of primer and template, the observed discrepancy reflected the degree of nucleotide divergence among HCV strains in the four plasmas.

Two sequences of 290 nucleotides (nt 3759–4048) and 354 nucleotides (nt 6427–6780)

TABLE 4 Two-by-two comparison of nucleotide and amino acid sequences within the three regions among the prototype HCV and HCV strains in the four plasmas[a].

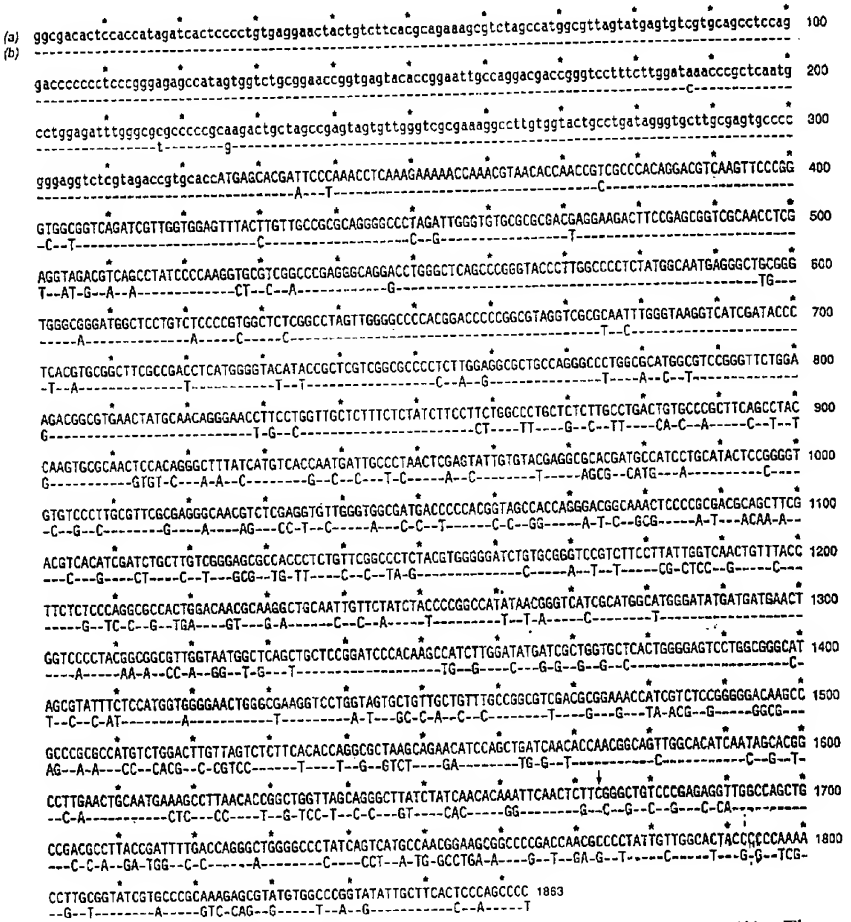
HCV strains in	Percentage divergence of nucleotides (amino acids)			
	Prototype HCV	HC-J1 strain	HC-J2 strain	HC-J3 strain
HC-J1	(*a*) 95.8 (93.9)	—	—	—
	(*b*) 95.2 (97.6)	—	—	—
	(*c*) 96.8 (98.1)	—	—	—
HC-J2	(*a*) 71.6 (78.8)	(*a*) 72.6 (78.8)	—	—
	(*b*) 77.2 (83.3)	(*b*) 75.6 (84.5)	—	—
	(*c*) 78.3 (82.1)	(*c*) 78.0 (82.1)	—	—
HC-J3	(*a*) 75.8 (81.8)	(*a*) 76.8 (84.8)	(*a*) 90.5 (87.9)	—
	(*b*) 77.2 (82.1)	(*b*) 74.4 (83.3)	(*b*) 90.8 (95.2)	—
	(*c*) 78.0 (83.0)	(*c*) 78.3 (83.0)	(*c*) 90.1 (91.5)	—
HC-J4	(*a*) NT[b]	(*a*) NT[b]	(*a*) NT[b]	(*a*) NT[b]
	(*b*) 74.8 (83.3)	(*b*) 72.0 (82.1)	(*b*) 91.6 (95.2)	(*b*) 92.8 (95.2)
	(*c*) 75.8 (82.1)	(*c*) 74.8 (82.1)	(*c*) 88.2 (86.8)	(*c*) 91.7 (93.4)

[a]: The prototype HCV (European Patent Applicaton #88310922.5) and HCV strains propagated from the four plasmas were compared against each other in the three regions and the results are given in percentage divergence. The regions in comparison were: (*a*), 95 nucleotides spanning nt 173–267; (*b*), 250 nucleotides spanning nt 3779–4028; (*c*), 314 nucleotides spanning nt 6447–6760. Nucleotides were numbered in accordance with the documented sequence of prototype HCV. Only nucleotides not overlapping with primers on both ends were compared.

[b]: Not tested because the (*a*) sequence was not displayed by the HCV strain in HC-J4.

were amplified on HCV strains in all four plasmas. Another sequences of 135 nucleotides (nt 153–287) was amplified on HCV strains in three plasmas (HC-J1, HC-J2 and HC-J3). The three regions of HCV genomes in the four plasmas, not overlapping with primer sequences at the both ends of PCR products, were subjected to two-by-two comparison along with the prototype HCV (TABLE 4). HC-J1 strain was homologous to the prototype HCV in greater than 95% for both nucleotide and amino acid sequences, but only in 71.6–78.3% for nucleotide sequence and 78.8–83.3% for amino acid sequence to HCV strains in the other three plasmas. HCV strains in these three plasmas (H-J2, HC-J3

Fig. 2. The 5′-terminal sequences of HCV strains in HC-J1 and HC-J4.



The sequence of HCV strain in HC-J1 is given in (*a*), and that in HC-J4 in (*b*). The uninterrupted open reading frame is shown in capitals which starts at the ATG codon occupying nt 325–327. The sequence upstream of it, spanning nt 1–324, represents the 5′-noncoding region and is expressed in lower case letters. The arrow indicates the start of the disclosed sequence of prototype HCV (European Patent Application #88310922.5) at nt 1673 in the numbering system.

and HC-J4), however, were homologous to each other in 88.2–92.8% for nucleotide sequence and 86.8–95.2% for amino acid sequence. HCV strains in HC-J1 and HC-J4, which showed the highest degree of divergence, were selected for the determination of the 5′-terminal sequence of HCV.

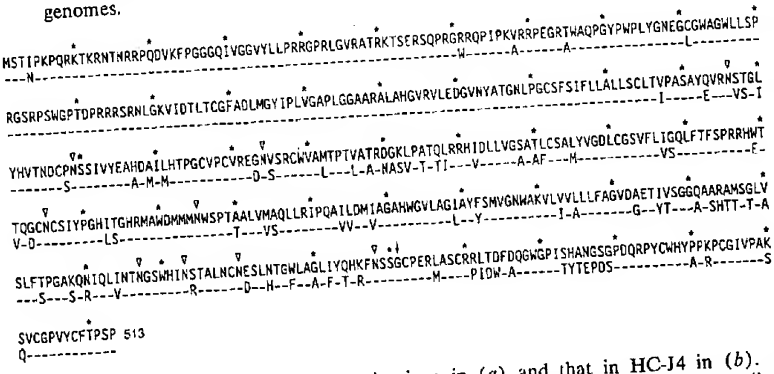
Nucleotide sequences of structural genes of two distinct HCV strains.

The strategy for sequencing is depicted in Fig. 1. The numbering of nucleotide sequence of HCV was started from the 5′-terminus commonly displayed by the two HCV strains in HC-J1 and HC-J4. At first, 1656 nucleotides representing nt 437–2092 were determined on the clone ϕ41; the clone was obtained from the cDNA library of HC-J1 primed with oligonucleotide #8. Based on this sequence, a new primer #25 (20-mer) was synthesized that represented nt 807–826. This primer, in turn, was applied to obtain four cDNA clones (ϕ60, 61, 66 and 75) from the cDNA library of HC-J4, on which the upstream sequence spanning nt 1–826 was determined.

The upstream sequence of HC-J1 strain was determined on three separate clones of PCR products primed with #44 and #25 (C2503, C2508 and C2510). The downstream sequence of 1163 nucleotides (nt 721–1883) of HC-J4 strain was determined on three independent clones of PCR products primed with #30 and #42 (C2821, C3173 and C3192). The primer #42 copied nt 1864–1883 of the HCV strain in HC-J3, and was used for this purpose because the HCV genome in HC-J4 did not anneal with any of primers #5, #6, #7 and #8 which copied the prototype HCV. Three clones of PCR products from either HC-J1 or HC-J4 plasma revealed divergence in only 1–2 nucleotides. The consensus sequence for 1863 nucleotides was deduced on each of the two HCV strains and compared against each other in Fig. 2.

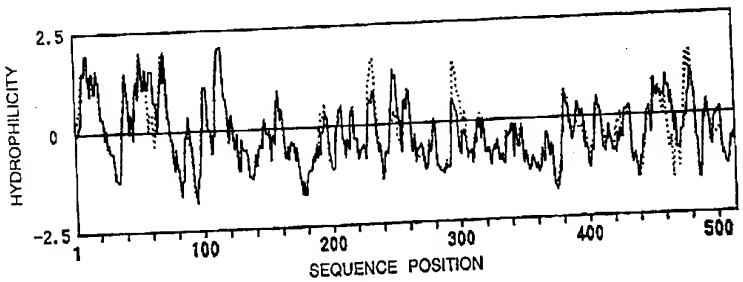
Both of them had a 5′-noncoding region of 324 nucleotides that were well conserved, differing only in 3 (0.9%). They had a long open reading frame starting at nt 325, which continued at nt 1673 to the documented prototype HCV (European Patent Application #88310922.5). Taken altogether, the HCV possessed a single long open reading frame,

Fig. 3 The predicted amino acid sequences of the products of structural genes of two HCV genomes.



The sequence of HCV genome in HC-J1 is given in (*a*) and that in HC-J4 in (*b*). Open triangles indicate potential N-glycosylation sites. The arrow indicates the start of the disclosed sequence of prototype HCV (European Patent Application #88310922.5) at the 451st amino acid in the open reading frame. Single letter abbreviations for amio acid residues are: A, alanine; R, arginine; N, asparagine; D, aspartic acid; C, cysteine; E, glutamic acid; Q, glutamine; G, glycine; H, histidine; I, isoleucine; L, leucine; K, lysine; M, methionine; F, phenylalanine; P, proline; S, serine; T, threonine; W, tryptophan; Y, tyrosine; V, valine.

Fig. 4 Profile of relative hydrophilicity of the products of structural genes of HCV strains.



Hydrophilicity score expressed by the six consecutive amino acid residues [17] is indicated along the 5′-terminal sequence of 513 amino acids. The solid line represents the HCV genome in HC-J1, and dotted line represents that in HC-J4.

bearing structural and non-structural genes, composed of at least 8658 nucleotides capable of encoding 2886 amino acid residues.

As regard the coding region of 1539 nucleotides (Fig. 2), HC-J1 and HC-J4 strains differed in 336 with a homology of 78.2%. The amino acid sequence of the two strains are compared in Fig. 3. They shared 83.0% of the encoded 513 residues. The two HCV strains had a homology of 97.2% within the region of amino acids 1–180. The homology decreased to 75.4% within the region of amino acids 181–513, however.

The profile of hydrophilicity, calculated by the method of Hopp and Woods [17], is depicted in Fig. 4. At least three areas of high local hydrophilicity were recognized in the upstream region corresponding to amino acids 1–120; this region was rich in basic amino acids (28/120 or 23%). In the downstream region spanning amino acids 121–513, the sequence tended to diverge accompanied by low hydrophilicity scores. Nine potential N-glycosylation sites were found restricted to the downstream region.

Discussion

The 5′-terminal sequence of the HCV genome, where structural genes are expected to reside, was determined for two distinct HCV strains propagated from human (HC-J1) and chimpanzee (HC-J4) plasmas provisioned to possess a high degree of divergence against each other. The nucleotide sequences of the two strains gave the opportunity to evaluate conserved and unconserved regions, and to postulate the locations of putative genes encoding capsid and envelope proteins of HCV. Approaches along this line have not been feasible, since only the nonstructural sequence corresponding to three-quarters of HCV genome is disclosed (European Patent Application #88310922.5).

The obtained 5′-terminal sequence of HCV showed no significant similarities with any of known sequences deposited in the latest release from a gene data bank (EMBL-GDB release 20.0 August, 1989). Within the coding region of 5′-terminal sequence of the two HCV strains, spanning nt 325–1863 and including 1539 nucleotides, there were no insertion or deletion of nucleotides and 336 (22.8%) were divergent; 159 were due to transversion mutation (purine to pyrimidine or *vice versa*), and 177 were due to transition mutation (purine to purine or pyrimidine to pyrimidine).

Among 336 point mutations, 157 (46.7%) resulted in amino acid substitutions. Mutations with amino acid changes were found in only 8 (17.0%) of 47 nucleotide alterations identified within codons 1–180, much less often than in 149 (51.6%) of 289 within codons 181–513. These results indicate that the

amino acid sequence encoded by the upstream region would be more conserved than that by the downstream region. The upstream region would encode the viral capsid protein, therefore, while the downstream region would encode the viral envelope protein. Mutations with amino acid changes cluster within the envelope genes of viruses, in general, reflecting efforts for survival on the part of virus against the selection of host immune responses.

Basic amino acids such as arginine and lysine prevailed in the upstream region (codons 1–120), accounting for 23%, while they were not so common in the downstream region (codon 121–513) and found only in 7.5%. Basic amino acids frequently occur in the capsid protein and implicated in the affinity for nucleic acids. Envelope proteins, in general, are associated with carbohydrate moieties. Nine potential N-glycosylation sites were found exclusively in the downstream region. Taken altogether, the upstream region of the 5′-terminal sequence would encode the capsid protein, while the downstream region would encode the envelope protein of HCV.

Identification of the putative capsid and envelope genes would enable their expression for developing immunoassays and vaccines. Estimated hydrophilic regions in the coded proteins (Fig. 4) would be candidates for antigenic epitopes, and synthetic oligopeptides mimicking their sequences would also be applied for these purposes.

In the 5′-terminal sequence, only the sequence of nt 1673–1863 (191 nucleotides) was available for comparison of the prototype HCV with the two HCV strains reported here. Within this nucleotide sequence, the prototype HCV was homologous in 92.7% to HC-J1 and in only 70.2% to HC-J4; the homology between HC-J1 and HC-J4 was 68.1%. Within a total of 564 nucleotides representing two separate regions encoding parts of nonstructural proteins (*b* and *c* in TABLE 4), HC-J2, HC-J3 and HC-J4 were close to each other but different to the prototype HCV or HC-J1 strain. At least two HCV genomes in Japan have been partially sequenced. One sequence [18], available for comparison in 160 nucleotides (nt 6601–6760), is homologous to HC-J1 strain in only 79.4%, while in 91.3–91.9% to HC-J2, HC-J3 and HC-J4 strains. The sequence of the other [19] does not overlap with any of the three regions in comparison. These results would indicate that HCV genomes can be classified into at least two groups based on the homology in the nucleotide sequence. More groups will emerge as sequence data accumulate, because HCV is an RNA virus that is liable to mutations.

It would be worthy of note that the 5′-noncoding region of 324 nucleotides was highly conserved in HC-J1 and HC-J4 strains, differing only in 3 (1%); HC-J2 and HC-J3 strains possessed the same 5′-noncoding region sequence as HC-J4 strain (unpublished observations). The 5′-noncoding region is well preserved by reported flaviviruses [20]. The terminus of the 5′-noncoding region of HCV, however, is based on the analysis on a single cDNA clone of HC-J4 strain (ϕ75).

Additional nucleotides, if any, on the 5′-noncoding region of the HCV genome would be revealed by sequencing cDNA clones of HCV strains in an extended series of plasmas; oligonucleotide primers copying the 5′-noncoding region would be very instrumental in such an undertaking. The 5′-noncoding sequence of at least 324 nucleotides is rather long, in comparison with 95 nucleotides in Japanese encephalitis virus [21], 96 in dengue type 2 virus [20] as well as West Nile fever virus [22] and 118 in yellow fever virus [23].

As a sensitive means to detect HCV in sera, PCR on cDNA with primers copied from the reported nonstructural sequence is proposed, which detected HCV in some sera without detectable anti-HCV [24]. The conservation of the 5′-noncoding region would be taken advantage of in designing primers to this end.

The taxonomy of HCV is still obscure. Based on chimpanzee transmission studies, the infectious agent of NANB hepatitis has been proposed as a virus of the Togaviridae family with a small diameter of 30–60 nm and associated with lipids [25]. The reported

amino acid sequence of the prototype HCV resembles the nonstructural number 3 (NS3) sequence of dengue type 2 virus in the flavivirus family [26]. The 5'-noncoding region of flaviviruses, however, is much shorter than that of HCV [20–23].

By means of computer-assisted protein analysis, a significant similarity of the sequence of the prototype HCV is reported, not only to flaviviruses and pestiviruses, but also to the members of the picornavirus-like and alphavirus-like plant virus supergroups [26]. Attempts along this line will gain in the precision, armored with the 5'-terminal sequence, which, in concert with the sequence of prototype HCV (European Patent Application #88310922.5), will cover most of the genome that is estimated at 10 kilobases in size.

References

[1] **Gocke, D. J.:** A prospective study of post-transfusion hepatitis. The role of Australia antigen. *J. Am. Med. Assoc.*, **219**, 1165–1170 (1972).

[2] **Alter, H. J., Purcell, R. H., Holland, P. V., Feinstone, S. M., Marrow, A. G.** and **Moritsugu, Y.:** Clinical and serological analysis of transfusion-associated hepatitis. *Lancet*, ii, 838–841 (1975).

[3] **Knodell, R. G., Conrad, M. E., Ginsberg, A. L., Bell, C. J.** and **Flannery, E. P.:** Efficacy of prophylactic gamma-globulin in preventing non-A, non-B post-transfusion hepatitis. *Lancet*, i, 557–564 (1976).

[4] **Feinstone, S. M., Kapikian, A. Z.** and **Purcell, R. H.:** Hepatitis A: detection by immune electron microscopy of a viruslike antigen associated with acute illness. *Science*, **182**, 1026–1028 (1973).

[5] **Prince, A. M., Brotman, B., Grady, G. F., Kuhns, W. J., Hazzi, C., Levine, R. W.** and **Millian, S. J.:** Long-incubation post-transfusion hepatitis without serological evidence of exposure to hepatitis-B virus. *Lancet*, ii, 241–246 (1974).

[6] **Tabor, E., Gerety, R. J., Drucker, J. A., Seeff, L. B., Hoofnagle, J. H., Jackson, D. R., April, M., Barker, L. F.** and **Pineda-Tamondong, G.:** Transmission of non-A, non-B hepatitis from man to chimpanzee. *Lancet*, i, 463–466 (1978).

[7] **Yoshizawa, H., Itoh, Y., Iwakiri, S., Kitajima, K., Tanaka, A., Nojiri, T., Miyakawa, Y.** and **Mayumi, M.:** Demonstration of two different types of non-A, non-B hepatitis by reinjection and cross-challenge studies in chimpanzees. *Gastroenterology*, **81**, 107–113 (1981).

[8] **Choo, Q.-L., Kuo, G., Weiner, A. J., Overby, L. R., Bradley, D. W.** and **Houghton, M.:** Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science*, **244**, 359–362 (1989).

[9] **Kuo, G., Choo, Q.-L., Alter, H. J., Gitnick, G. L., Redeker, A. G., Purcell, R. H., Miyamura, T., Dienstag, J. L., Alter, M. J., Stevens, C. E., Tegtmeier, G. E., Bonino, F., Colombo, M., Lee, W.-S., Kuo, C., Berger, K., Shuster, J. R., Overby, L. R., Bradley, D. W.** and **Houghton, M.:** An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis. *Science*, **244**, 362–364 (1989).

[10] **Esteban, J. I., Esteban, R., Viladomiu, L., López-Talavera, J. C., González, A., Hernández, J. M., Roget, M., Vargas, V., Genesca, J., Buti, M., Guardia, J., Houghton, M., Choo, Q.-L.** and **Kuo, G.:** Hepatitis C virus antibodies among risk groups in Spain. *Lancet*, ii, 294–297 (1989).

[11] **van der Poel, C. L., Reesink, H. W., Lelie, P. N., Leentvaar-Kuypers, A., Choo, Q.-L., Kuo, G.** and **Houghton, M.:** Anti-hepatitis C antibodies and non-A, non-B post-transfusion hepatitis in the Netherlands. *Lancet*, ii, 297–298 (1989).

[12] **Alter, H. J., Purcell, R. H., Shih, J. W., Melpolder, J. C., Houghton, M., Choo, Q.-L.** and **Kuo, G.:** Detection of antibody to hepatitis C virus in prospectively followed transfusion recipients with acute and chronic non-A, non-B hepatitis. *N. Engl. J. Med.*, **321**, 1494–1500 (1989).

[13] **Miyamura, T., Saito, I., Katayama, T., Kikuchi, S., Tateda, A., Houghton, M., Choo, Q.-L.** and **Kuo, G.:** Detection of antibody against antigen expressed by molecularly cloned hepatitis C virus cDNA: application to diagnosis and blood screening for posttransfusion hepatitis. *Proc. Natl. Acad. Sci. U.S.A.*, **87**, 983–987 (1990).

[14] **Stevens, C. E., Taylor, P. E., Pindyck, J., Choo, Q.-L., Bradley, D. W., Kuo, G.** and **Houghton, M.:** Epidemiology of hepatitis C virus. A preliminary study in volunteer blood donors. *J. Am. Med. Assoc.*, **263**, 49–53 (1990).

[15] **Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B.** and **Erlich, H. A.:** Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, **239**, 487–491 (1988).

[16] **Sanger, F., Nicklen, S.** and **Coulson, A. R.:** DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 5463–5467 (1977).

[17] **Hopp, T. P.** and **Woods, K. R.:** Prediction of protein antigenic determinants from amino acid sequences. *Proc. Natl. Acad. Sci. U.S.A.*, **78**, 3824–3828 (1981).

[18] **Kato, N., Ohkoshi, S.** and **Shimotohno, K.:** Japanese isolates of non-A, non-B hepatitis